Journal of General Education and Humanities
Vol. 5, No. 4, August 2026, pp. 5739 – 5753, https://doi.org/10.58421/gehu.v5i4.1933
ISSN 2963-7147

Testing the Effect of the Smart Counseling System AI Support Vector Machine Intervention on Reducing Teen Misbehavior

Risydah Fadilah[1], Sari Atika Parinduri[2], Nuraini Kemalasari Istiqomah[3]
[1]Universitas Medan Area, Medan, Indonesia
[2]Sekolah Tinggi Agama Islam Tebing Tinggi Deli, Sumatera Utara, Medan Indonesia
[3]Universitas Panca Budi, Medan, Indonesia

Article Info

Article history:

Received 2026-07-27
Revised 2026-08-21
Accepted 2026-08-27

Keywords:

Adolescent Behavioral Difficulties
Artificial Intelligence
Guidance and Counseling
Smart Counseling System
Strengths and Difficulties Questionnaire
Support Vector Machine

ABSTRACT

The study aimed to evaluate the Artificial Intelligence-assisted Smart Counseling System in supporting the detection and classification of adolescent behavioral difficulties and to identify changes in students' behavioral conditions following the intervention. The study used a quantitative pretest–posttest control group design with 666 junior high school students in Medan City: 333 in the experimental group and 333 in the control group. Data were collected using the Strengths and Difficulties Questionnaire and analyzed by comparing pretest and posttest scores, as well as classifying them using the Support Vector Machine method. The results showed that the experimental group's mean Strengths and Difficulties Questionnaire score decreased from 21.00 (standard deviation = 12.469) at the pretest to 18.05 (standard deviation = 8.975) at the posttest. However, the analysis indicated that the difference between the groups at the posttest was not statistically significant, $F(2,663) = 1.578$, $p = 0.207$. Therefore, the evidence was insufficient to conclude that the Smart Counseling System was significantly more effective than the control condition. In contrast, the Support Vector Machine analysis showed strong classification performance, with accuracy increasing from 98.50% using the original features to 99.25% after feature engineering for classifying students into the Normal, Borderline, and Abnormal categories. These findings indicate that the Smart Counseling System has potential as a supporting system for the early detection and classification of adolescent behavioral difficulties; however, its effectiveness as an intervention requires further investigation through analyses of time-by-group interactions, effect sizes, and more comprehensive model validation. Therefore, the Smart Counseling System is better positioned as a decision-support tool for counselors rather than a replacement for counselors or a diagnostic instrument.

This is an open-access article under the CC BY-SA license.

Corresponding Author:

Risydah Fadilah
Fakultas Psikologi, Universitas Medan Area, Medan, Indonesia
Email: risydah@staff.uma.ac.id

1. INTRODUCTION

The teenage years are a stage of development marked by complex physical, psychological, social, and moral changes. At this stage, teens are expected to build their self-identity, develop independence, adjust to their social environment, and manage their emotions and behavior [1, 2]. However, this development process does not always run smoothly, so various behavioral problems can arise, such as breaking school rules, conflicts with peers, aggression, skipping classes, impulsive behavior, and actions that go against social norms. These behavioral difficulties not only affect teenagers' psychosocial development but can also hinder learning and disrupt the creation of a safe, conducive school environment [3, 4]. Therefore, guidance and counseling services are needed to identify and address behavioral problems systematically.

Urban contexts like Medan City bring increasingly complex social dynamics for adolescent development. Intense social interactions, digital technology development, social media use, peer influence, and changes in relationships among teens, families, and schools can all contribute to problematic behavior. Research documents that form the basis of this study show that juvenile delinquency is related to various factors, including low self-control, peer influence, parenting patterns, weak parental monitoring, and the social environment. This situation highlights the importance of strengthening guidance and counseling services in schools so they can detect students facing behavioral difficulties early and provide interventions according to their needs.

The development of artificial Intelligence (AI) offers new opportunities to improve guidance and counseling services through the development of a Smart Counseling System (SCS). This system can help collect, process, and classify student behavior data so counselors have more systematic information for determining service needs [5]. In the research that forms the basis of the study, SCS uses data from the Strengths and Difficulties Questionnaire (SDQ), which includes 25 items related to emotional, behavioral, social aspects, hyperactivity, and prosocial behavior. These data are classified into Normal, Borderline, and Abnormal categories using a Support Vector Machine (SVM) algorithm [6, 7]. With this approach, AI can serve as a tool to support counselors in risk mapping and early detection, not as a replacement for professional counselors.

Evaluation of AI-assisted counseling systems should be based not only on the algorithm's ability to classify, but also on changes in students' behavior after receiving the intervention. Therefore, a controlled pretest–posttest design is important to compare students' conditions before and after implementing the SCS, as well as to compare them with a group that did not receive the intervention. The basic study involved 666 middle school students divided into an experimental group and a control group, with 333 students in each group. In addition, SVM analysis showed that student data fell into 95 Normal, 174 Borderline, and 397 Abnormal categories, making class imbalance another important aspect in evaluating the model's performance.

Research on the use of Artificial Intelligence (AI) in adolescent mental health has progressed rapidly. Zhang, through a systematic scoping review of 88 studies, found that AI is most commonly used for diagnosis, followed by monitoring and evaluation, therapy, and prognosis [8]. The study also showed that machine learning is the most commonly used AI method, but most research still focuses on diagnostic functions, so the intervention aspect and evaluating changes in adolescents' conditions have not been explored in depth. Next, Komisarow & Hemelt conducted an experimental study with 56 teenagers and compared AI-based counseling, human counselors, and human-AI collaboration through a pretest–posttest design [9]. The results show that chat-based counseling can reduce negative feelings and improve mood, but AI alone is less effective than human counseling. Zhang et al.'s research, through the development of BetterMood, also shows the potential of AI counseling services that mimic human interaction to overcome the limited access to face-to-face counseling for teens and young adults [10].

Although previous research has shown the potential of AI in supporting adolescent mental health, a research gap remains in integrating counseling interventions, evaluating behavior changes, and classifying risk using machine learning in a single study design. Liu et al.'s research shows that most AI studies on adolescent mental health focus on diagnosis and prediction [11]. Meanwhile, Turki et al. focused more on comparing the effectiveness of different forms of counseling on mental health indicators [12]. Meanwhile, Razi et al., in a scoping review of 24 studies, found that AI has potential in risk prediction, classification, service personalization, and improving access to adolescent mental health services. However, issues remain, including algorithmic errors, reliance on self-reported data, privacy concerns, transparency, and limited validation in real-world conditions [13]. Thus, research is still needed to examine whether AI systems that can classify behavioral difficulties can also produce measurable changes after counseling interventions, especially in the context of education in Indonesia and the city of Medan.

The novelty of the study "Evaluating an AI-Assisted Smart Counseling System for Adolescent Behavioral Problems: A Controlled Pretest–Posttest Study and SVM Classification Analysis" lies in integrating three components: counseling intervention evaluation, a controlled pretest–posttest design, and Support Vector Machine (SVM) classification. Unlike previous studies that often separate AI classification from counseling effectiveness, this study uses SVM to map students' behavioral problem levels and then assesses changes in students' conditions after interventions by comparing experimental and control groups. This approach creates a "classification–intervention–outcome" framework, so AI is evaluated not only based on technical accuracy but also on its relevance to changes in student behavior. In addition, this study adds contextual novelty by applying it to school students in Medan City, generating empirical evidence on implementing an AI-based Smart Counseling System in the Indonesian education setting. This approach aligns with recent findings that research on AI in adolescent mental health still requires more validation in real-world contexts, stronger longitudinal/experimental designs, and system development that still places professionals as the main decision-makers.

2. METHOD

This study uses a quantitative, experimental approach to evaluate the application of the Smart Counseling System (SCS), assisted by Artificial Intelligence (AI), in reducing behavioral difficulties among adolescents in junior high schools in Medan City. The research

 https://doi.org/10.58421/gehu.v5i4.1933

uses a pretest–posttest control group design, in which participants' conditions are measured before and after the intervention in both an experimental and a control group. This design allows researchers to observe changes within groups and compare changes between the group that received the intervention and the group that did not. In a controlled experimental design, measuring participants before the intervention is important to identify their initial conditions so researchers can evaluate changes after treatment objectively [14]. Based on the research design, the participants were 666 junior high school students in grades 7–9, with 333 in the experimental group and 333 in the control group.

The research design is laid out operationally in Table 1. The experimental group receives an initial measurement (pretest), then an intervention in the form of an AI-assisted Smart Counseling System, and finally a final measurement (posttest). The control group also gets pretests and posttests but does not receive the SCS intervention during the study. Using a control group lets us see whether changes in the experimental group's scores are due to the intervention rather than natural changes or other factors in the control group. If the 666 participants are assigned to the experimental and control groups completely at random, this design is called a randomized pretest–posttest control group design. On the other hand, if participants come from already formed classes or schools without individual randomization, the more accurate term is quasi-experimental pretest–posttest control group design. [15]. This distinction matters so the methodological design matches the procedures actually carried out in the research.

Table 1. Research Design and Procedure

Group	Pretest (O_1)	Treatment	Posttest (O_2)
Experimental (n = 333)	SDQ assessment	Artificial Intelligence-assisted *Smart Counseling System*	SDQ assessment
Control (n = 333)	SDQ assessment	Did not receive the Smart Counseling System	SDQ assessment

The independent variable was the Artificial Intelligence-assisted Smart Counseling System intervention, and the dependent variable was adolescent behavioral difficulties. Behavioral difficulties were measured using the *Strengths and Difficulties Questionnaire* (SDQ). The SDQ is an instrument consisting of 25 items covering five dimensions: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. Each dimension comprises five items and uses a three-point response scale. The SDQ can screen for psychosocial problems among adolescents, but its factor structure and each subscale's reliability should be interpreted in line with the target population's characteristics. Furthermore, research involving Indonesian adolescents has shown that the self-report version of the SDQ can identify behavioral problems. However, screening results should be interpreted as an initial screening measure rather than as a substitute for clinical or professional assessment [16].

The research data were then processed to support two analysis goals: evaluating behavior changes and AI-based classification. In the first stage, the SDQ responses were checked for completeness, coded, and scored, which resulted in individual participants' behavior difficulty scores. In the second stage, the SDQ data were used as input for a Support Vector Machine (SVM) model. The SVM classified participants into Normal, Borderline, and Abnormal categories based on their response patterns. SVM is relevant because it can handle classification in high-dimensional feature spaces and supports the development of data-driven decision support systems [17].

Research on Support Vector Machines shows that the algorithm can classify data effectively, especially when the data has a relatively large number of features. SVM's ability to find the optimal separating boundary (hyperplane) makes it suitable for various classification problems and for developing data-driven decision support systems. In this study, SVM is used to classify patterns of student behavioral difficulties based on responses to the Strengths and Difficulties Questionnaire, providing supporting information for early detection and guidance and counseling services [18].

During the intervention stage, the experimental group received the AI-assisted Smart Counseling System program, which was designed as a support system for guidance and counseling services. The system provides behavior classification based on SDQ data so counselors can get an initial picture of students' conditions and determine the need for follow-up. In this study, AI is not positioned as a replacement for counselors, but as a supporting tool for early detection and professional decision-making. After the intervention, both the experimental and control groups completed the SDQ again as a posttest. The changes in scores between the pretest and posttest were then analyzed to determine whether behavior difficulties changed differently between the two groups [19].

The effectiveness of the intervention was analyzed using Mixed ANOVA, with measurement time (pretest and posttest) as the within-subject factor and research groups (experiment and control) as the between-subject factor. This analysis examined three main effects: changes in scores over time, differences in scores between groups, and the time-by-group interaction. The interaction effect is the main focus because it shows whether the change in behavior difficulty in the experimental group differs significantly from that in the control group. Before doing the Mixed ANOVA, the data were checked through appropriate assumption tests, including normality and homogeneity of variance. If the parametric assumptions were not met, alternative analyses could be used depending on the characteristics of the data [20].

Meanwhile, the SVM model evaluation was done using a confusion matrix, accuracy, precision, recall, and F1-score. These indicators are necessary so the model's performance is judged not only by the total number of correct predictions, but also by its ability to recognize each category of behavioral difficulties. This becomes especially important if the data are imbalanced across the Normal, Borderline, and Abnormal categories. Research on SVM classification with high-dimensional data shows that feature selection can help reduce complexity, remove redundant features, and improve prediction performance [21]. Thus, the SVM analysis in this study aims not only to produce a high-accuracy model but also to provide more consistent classification information to support counseling services.

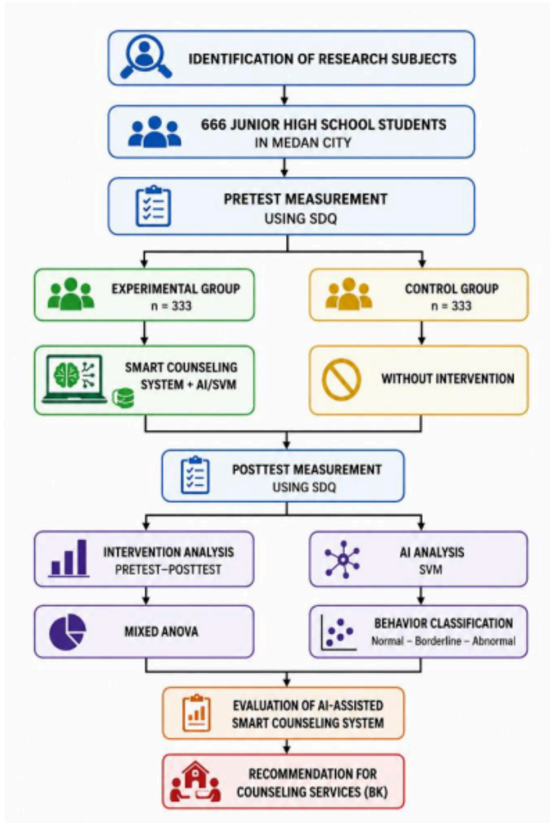

Figure 1. Flow of Assisted SCS Research Methodology

The image shows that the research does not stop at developing the SVM algorithm, but connects behavioral data → AI classification → counseling intervention → behavior change. This framework strengthens the research's position because the success of SCS is evaluated from two perspectives at once, namely the effectiveness of the intervention statistically and the performance of the classification computationally. The SDQ also has empirical support in adolescent populations, including validity evidence for screening psychosocial problems [22].

Table 2. Variables, Indicators, and Analysis Techniques

Variable	Indicator	Instrument/Method	Analysis
SCS-AI Intervention	Implementation of the Smart Counseling System	*Smart Counseling System*	Descriptive analysis
Behavioral difficulties	Emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior	25-item Strengths and Difficulties Questionnaire (SDQ)	Mixed ANOVA
Behavioral classification	Normal, Borderline, and Abnormal	Support Vector Machine (SVM)	Accuracy, precision, recall, and F1-score
Behavioral changes	Differences between pretest and posttest scores	Strengths and Difficulties Questionnaire (SDQ)	Mixed ANOVA
Model performance	Classification accuracy	SVM and confusion matrix	Model evaluation

With this design, the study produced two types of empirical evidence. The first is evidence of changes in students' behavioral difficulties after the SCS intervention, based on a comparison of the experimental and control groups. The second is evidence about AI's ability to classify patterns of behavioral difficulties based on SDQ data. The integration of these two pieces of evidence serves as the basis for evaluating the feasibility of SCS as a data-driven guidance and counseling support system in schools in Medan.

3. RESULTS AND DISCUSSION

3.1. Results

Based on the results presented in the attached manuscript, the Results section needs revision to ensure consistency with the correct statistical outputs. In particular, the ANOVA results showed a *p*-value of 0.207 for the posttest; therefore, it would be inappropriate to conclude that the intervention was effective based on differences between groups. Furthermore, Table 6 should not be described as a Wilcoxon test because the *F* column is inconsistent with the output of a Wilcoxon analysis. The *Z* value, positive ranks, negative ranks, and effect size will not be fabricated because these values are not available in the manuscript. Therefore, the following presentation provides a more statistically accurate and academically appropriate interpretation.

3.1.1. Test of Homogeneity

Before analyzing differences in behavioral difficulty scores, we conducted a test of variance homogeneity using Levene's Test. The results showed a significance value of 0.901 for the *pretest* scores and 0.975 for the *posttest* scores. Both values were greater than 0.05, indicating that the variances across groups met the assumption of homogeneity. Therefore, the data met one of the assumptions required for subsequent parametric analysis.

Table 3. Results of the Homogeneity of Variance Test

Measurement	Levene Statistic	df1	df2	Sig.	Interpretation
Pretest SDQ	0.105	2	663	0.901	Homogeneous
Posttest SDQ	0.026	2	663	0.975	Homogeneous

3.1.2. Comparison of Pretest and Posttest Scores Across Groups

The ANOVA results indicated no statistically significant difference between groups at the *pretest* stage, with $F(2, 663) = 0.909$; $p = 0.404$. This finding indicates that participants' initial conditions across the compared groups were statistically comparable. At the *posttest* stage, the analysis yielded $F(2, 663) = 1.578$; $p = 0.207$. Since the *p*-value was greater than 0.05, there was no statistically significant difference between groups in the *posttest* measurement. Therefore, based on the available between-group analysis, this study does not provide sufficient statistical evidence that participants who received the Smart Counseling System (SCS) had significantly different posttest scores compared with the comparison groups.

 https://doi.org/10.58421/gehu.v5i4.1933

Table 4. Results of the Analysis of Differences in SDQ Scores Across Groups

Measurement	F	df	Sig. (p)	Interpretation
Pretest SDQ	0.909	2, 663	0.404	Not significant
Posttest SDQ	1.578	2, 663	0.207	Not significant

These findings are important because the p-value of 0.207 at the *posttest* stage cannot be used as a basis for concluding that the SCS intervention was effective. Descriptive changes or observed score differences alone are insufficient to establish a treatment effect. Therefore, conclusions about the intervention's effectiveness should be based on an analysis that directly examines the interaction between measurement time (pretest–posttest) and group (experimental–control), as expected in a Mixed ANOVA. Since the interaction output is not available in the provided manuscript, the effectiveness of the intervention cannot be definitively established based solely on the available ANOVA results.

3.1.3. Changes in Scores in the Experimental Group

Descriptively, the experimental group, consisting of 333 students, showed a change in scores from pretest to *posttest*. The mean *pretest* score was 21.00 (SD = 12.469), whereas the mean *posttest* score was 18.05 (SD = 8.975). Thus, the mean score decreased by 2.95 points following implementation of the SCS program. This descriptive finding indicates a tendency toward reduced behavioral difficulty scores among students in the experimental group.

Table 5. Descriptive Statistics of the Experimental Group

Measurement	N	Mean	SD	Mean Change
Pretest	333	21.00	12.469	–
Posttest	333	18.05	8.975	−2.95

However, this decrease in the mean score should not be directly interpreted as evidence of intervention effectiveness. Changes within a single group may be attributable to individual variation, natural changes during the study period, students' experiences outside the intervention, or other uncontrolled factors. To establish that the observed decrease was attributable to the SCS intervention, changes from *pretest* to *posttest* should be evaluated in comparison with the control group, particularly through the time × group interaction effect in a Mixed ANOVA. Since this output is not available in the current manuscript, the appropriate interpretation is that the experimental group demonstrated a descriptive decrease in scores, rather than that the SCS intervention has been inferentially proven effective.

3.1.4. Analysis of the Smart Counseling System Classification Using SVM

In addition to analyzing changes in behavioral difficulty scores, the study conducted classification using a Support Vector Machine (SVM) based on *Strengths and Difficulties Questionnaire* (SDQ) data. The dataset consisted of three categories: *Normal*, *Borderline*, and *Abnormal*. The data distribution indicated that the *Abnormal* category was the largest, comprising 397 observations, followed by the *Borderline* category with 174 observations and the *Normal* category with 95 observations. This distribution indicates class imbalance, which should be considered when interpreting the classification model's performance.

Table 6. Distribution of the SVM Classification Dataset

Category	Number	Percentage
Normal	95	14.25%
Borderline	174	26.09%
Abnormal	397	59.61%
Total	666	100%

The model development results showed that the linear SVM with feature engineering achieved an accuracy of 0.9925, an improvement over the 0.9850 accuracy obtained using the original features. Feature engineering integrated SDQ subscales and interaction terms, enabling the model to capture more complex patterns in the data. The manuscript also indicates that this improvement was associated with reduced misclassification within the *Borderline* category.

Table 7. SVM Model Performance

Model	Features	Accuracy
Linear SVM	Original features	0.9850
Linear SVM	Engineered features	**0.9925**

These results indicate that the SVM demonstrated very high classification performance on the SDQ dataset used in the study. However, interpret the high accuracy cautiously because the dataset has an imbalanced class distribution. Therefore, model evaluation should ideally be complemented by a confusion matrix, precision, recall, and F1-score for each class, particularly to determine whether the predominance of the Abnormal category substantially influences the high accuracy. The current manuscript does not provide all of these values; therefore, they should not be added without the corresponding original analytical outputs.

Overall, the findings indicate two main results. First, the experimental group's mean score decreased from 21.00 to 18.05, indicating a descriptive change following implementation of the SCS program. However, the between-group comparison at *posttest* yielded F = 1.578 and $p = 0.207$, indicating that the observed difference was not statistically significant. Therefore, based on the available data, the study cannot conclude that the Smart Counseling System significantly reduced behavioral difficulties compared with the control group. Second, from an artificial intelligence perspective, the SVM model demonstrated very high classification performance, achieving 99.25% accuracy after feature engineering, indicating the potential of SCS as a data-driven behavioral classification support tool.

Thus, the findings suggest that SCS has potential as a support system for detecting and classifying behavioral difficulties among adolescents. However, based on the available outputs, its effectiveness as an intervention to reduce behavioral difficulties has not yet been statistically established. To establish the intervention effect, subsequent analyses should use a complete Mixed ANOVA, reporting the F value, p-value, and effect size for the time ×

group interaction. If the researchers intend to use the Wilcoxon test, Table 6 should instead be recalculated using the Wilcoxon Signed-Rank Test and reported with positive ranks, negative ranks, ties, Z, p-value, and effect size. These values should not be inferred from the current table because the original Wilcoxon output is not available.

3.2. Discussion

The research results show that there was no significant difference between groups at the pretest stage, with values of (F(2,663)=0.909) and (p=0.404), indicating that the initial condition of the participants was relatively comparable. The homogeneity test also showed that the pretest and posttest variances met the assumption of homogeneity, with significance values of 0.901 and 0.975, respectively. These conditions provide an adequate methodological basis for comparing changes in participants' conditions after the intervention. However, at the posttest stage, the differences between groups were not significant, with (F(2,663)=1.578) and (p=0.207). These findings suggest that although scores changed within the experimental groups, the available statistical evidence is not sufficient to claim that the Smart Counseling System (SCS) is significantly more effective than the control group. Methodologically, conclusions about effectiveness should be based on the time × group interaction in a Mixed ANOVA, not just on posttest score comparisons. So, interpret the results carefully as evidence of a trend of change, not as definitive proof of the intervention's effectiveness.

Descriptively, the experimental group's average score decreased from 21.00 (SD = 12.469) on the pretest to 18.05 (SD = 8.975) on the posttest, a drop of 2.95 points. This pattern indicates that after participating in the SCS program, behavioral difficulty scores shifted downward. The findings are relevant to the literature on technology-based interventions for adolescents. Cen, through a systematic scoping review of 88 studies, found that AI has been used in various aspects of adolescent mental health, including diagnosis, monitoring, intervention planning, and therapy [23]. However, the study also showed that AI applications remain much more focused on diagnosis and assessment than on interventions, with 78 studies on diagnosis, 19 on monitoring/evaluation, 10 on treatment, and 6 on prognosis. Therefore, the trend of score decreases in this study is promising, but it also highlights the need for stronger experimental research to confirm whether these changes are caused by the SCS intervention.

Research findings can also be compared with the development of research on AI as a medium for psychological intervention. Recent literature shows that AI use in adolescent mental health is no longer limited to screening processes, but is shifting toward chat-based counseling, conversational agents, digital therapeutics, and decision support systems. A systematic review on AI in child and adolescent interventions identified 38 studies using AI as conversational agents, game-based platforms, or robotic systems for psychotherapy and developmental support. However, the presence of AI as an intervention medium does not automatically guarantee therapeutic success. Xuanfeng, in a scoping review on AI-based mental health support and therapy for adolescents, showed that AI has potential in risk prediction, classification, personalized support, and improving service access. However, challenges remain aroundeal-world validation, algorithmic errors, privacy, transparency, and reliance on self-reported data [24]. This study's findings support that argument: SCS shows strong technological potential, but the effectiveness of its interventions can't be concluded from the drop in scores in the experimental group alone.

From a measurement perspective, using the Strengths and Difficulties Questionnaire (SDQ) in this study has a strong empirical basis because the SDQ is a 25-item instrument designed to measure various aspects of psychosocial adjustment and behavioral difficulties in children and adolescents. Kim & Song's research on more than 10,000 children showed that the five-factor structure of the SDQ emotional symptoms, conduct problems, hyperactivity, peer problems, and prosocial behavior has psychometric support and that high scores are linked to an increased likelihood of independently identified psychological disorders. [25]. Li also showed that combined profiles of self-assessment and parent assessments using the SDQ can provide useful information for predicting healthcare usage and psychiatric diagnoses [26]. Timmons et al.'s research shows that the SDQ can screen for disruptive behavior in at-risk teenagers, so using it in studies of adolescent behavioral problems has empirical relevance [27]. Thus, choosing the SDQ as a data source for the SCS is justified, especially since the instrument captures behavioral and psychosocial dimensions relevant to early detection needs in schools.

Even so, this study's results also need to consider the debate about the SDQ's validity and reliability in adolescent groups. Tripathi & Alam, using data from 23,980 Finnish adolescents, found that the SDQ structure works better in early teens, but some models fit the mid-teens poorly, and a few subscales have low reliability [28]. The author even warns that using the total score or SDQ subscales requires careful consideration. In contrast, Nassif et al. support the SDQ's construct validity through a multitrait-multimethod approach, and a study of 1,489 Australian teenagers also supports the SDQ's internal structure and its relationships with variables such as loneliness, belonging, depression, and bullying [29]. These differences suggest that the SDQ's quality can be influenced by age, population, culture, and reporting methods. Therefore, the SCS classification results based on the SDQ in this study should be seen as a screening and early-detection tool, not a psychological diagnosis or a definitive determination of teen delinquency.

The strongest findings in this study actually came from the SVM classification analysis. The dataset consisted of 95 students in the Normal category, 174 in the Borderline category, and 397 in the Abnormal category. Thus, the Abnormal category dominates the dataset, so class imbalance must be considered when interpreting the model. The linear SVM model with the original features achieved an accuracy of 0.9850, while after feature engineering, the accuracy increased to 0.9925. These results show that SVM can recognize the SDQ response patterns used in the study with high accuracy. These findings align with systematic reviews on machine learning in mental health diagnosis prediction, which identify SVM as one of the algorithms frequently used alongside random forest, deep neural networks, and other methods [30]. However, the review also emphasized the need for larger, more diverse datasets, as well as longitudinal data so that models can capture the dynamics of mental health over time.

The increase in SVM accuracy after feature engineering also shows how important data representation is in AI systems. SDQ data includes not only individual scores but also relationships among emotional, behavioral, hyperactivity, peer, and prosocial dimensions. When this information is represented through more informative features, the model can pick up patterns that were not fully visible in the original features. The findings show an accuracy increase from 98.50% to 99.25%, indicating that feature engineering adds useful information for the model. However, this high accuracy should not be taken immediately as proof that the model is completely reliable. A recent meta-analysis of 188 studies involving 4,794,001 adolescent participants shows that AI models in mental health assessment can provide good discrimination, but their performance varies depending on the type of outcome, data source, and validation method; the AUC for internal validation ranges from 0.75–0.85, while external validation ranges from 0.76–0.96. This comparison highlights the importance of distinguishing accuracy on research datasets from the model's ability to generalize to new populations.

Class imbalance is another important aspect to discuss. Since the Abnormal category accounts for 397 of 666 data points, the model can achieve high accuracy if most predictions correctly identify the majority class. Therefore, SVM performance should be reported not only by accuracy, but also by precision, recall, F1-score, and a confusion matrix for each category. This approach lets researchers assess whether the model can truly identify the Normal and Borderline categories, which have smaller sample sizes, or whether it only excels at identifying the Abnormal category. This aligns with Yang et al.'s recommendation that developing AI in adolescent mental health requires transparency in model reporting, data management, analytical processes, and proper validation before applying it in practice [31]. So, an accuracy of 99.25% is promising, but it still needs to be complemented with class performance indicators and external validation to be considered a strong, generalizable model.

From a guidance and counseling service perspective, this research shows that SCS can shift the service model from a purely reactive approach to a more preventive, data-based one. In conventional practice, counselors often rely on teacher reports, observations, interviews, and students revealing their issues. SCS can help process data systematically so counselors can identify students showing certain risk patterns early. However, AI classification results should not be the sole basis for labeling students. The system should be positioned as an early-warning tool that helps counselors identify who needs further assessment, not as a tool to label a student as "troubled" or "naughty." This principle is becoming increasingly important because recent literature shows that AI in adolescent mental health still faces issues like ethics, algorithmic bias, privacy, transparency, and limited validation. So, the ideal SCS implementation is human-in-the-loop: the algorithm handles initial screening and classification, while professional counselors handle interpretation, counseling, and intervention decisions.

The research findings also have implications for how we understand juvenile delinquency. In the text, juvenile delinquency includes rule violations, truancy, uncontrolled behavior, fighting, and other forms of deviant behavior. However, the SDQ measures strengths and difficulties, not the full construct of juvenile delinquency. Therefore, using the term 'reduction in juvenile delinquency' needs to be done more carefully. A decrease in SDQ scores is more accurately described as a decrease in behavioral/psychosocial difficulties measured by the SDQ. Sidhu and Sidhu did show a relationship between the SDQ and disruptive behavior in at-risk adolescents, but that relationship does not mean that the SDQ is identical to all forms of juvenile delinquency [32]. This conceptual distinction matters so research conclusions do not go beyond the constructs actually measured.

Overall, this study produced two distinct but complementary contributions. First, on the intervention dimension, the experimental group's scores tended to drop from 21.00 to 18.05, but the posttest differences between groups were not significant ($p = 0.207$), so the available output does not yet statistically confirm the effectiveness of SCS as an intervention. Second, on the technology dimension, SVM showed very high classification performance (99.25% accuracy) after feature engineering, providing preliminary evidence that AI can help classify student behavioral difficulties. This pattern aligns with recent literature that positions AI as a promising technology for adolescent mental health assessment and support, while also emphasizing the need for external validation, experimental research, algorithm transparency, and professional oversight.

So, the empirical novelty of this study should not rest on claims that SCS has been proven to reduce juvenile delinquency, but on the effort to integrate counseling interventions, SDQ measurements, and SVM classification into a single AI-based guidance and counseling service system. This study shows that AI classification can be highly accurate even when evidence of intervention effectiveness is still limited. In fact, that difference becomes an important finding because it shows that an "accurate model" is not always the same as an "effective intervention." Going forward, research should use a full Mixed ANOVA by reporting the time × group interaction, effect size, confidence interval, and longitudinal analysis; the SVM model should be tested through cross-validation, external validation, a confusion matrix, precision, recall, and F1-score. With these improvements, SCS can evolve from a classification model into a more valid, transparent, adaptive, and responsible counseling support system for school guidance and counseling services.

4. CONCLUSION

Overall, this study indicates that the *Smart Counseling System* assisted by Artificial Intelligence has potential as a supporting system for the early detection and classification of behavioral difficulties among adolescents. However, it cannot yet be considered demonstrably effective as an intervention because the difference in posttest scores between the groups was not statistically significant ($F(2,663) = 1.578$; $p = 0.207$), although the experimental group's mean score decreased from 21.00 to 18.05. From a technological perspective, the Support Vector Machine showed strong classification performance, with accuracy increasing from 98.50% to 99.25% after feature engineering. Nevertheless, class imbalance and the incomplete reporting of precision, recall, F1-score, and confusion matrix indicators remain limitations that should be addressed. Therefore, the *Smart Counseling System* is better positioned as a tool for early detection and counselor decision support rather than a replacement for counselors or a diagnostic instrument. Future research should apply a complete Mixed Analysis of Variance procedure, including an analysis of the time × group interaction and effect size, strengthen Support Vector Machine validation through cross-validation and external validation, address class imbalance, and employ longitudinal designs and more diverse data sources to enable a more comprehensive evaluation of the system's effectiveness, accuracy, generalizability, and safety.

REFERENCES

[1] M. Lu, "Social support, resilience, and self-esteem protect against common mental health problems in early adolescence A nonrecursive analysis from a two-year longitudinal study," *Medicine (Baltimore),* vol. 100, no. 4, p. e24334, 2021.

[2] B. Kadiyala, R. P. Nippatla, S. Boyapati, C. Vasamsetty, S. K. Alavilli and Thanjaivadivel, "RETRACTED ARTICLE: Prevention and health care intervention of common injuries in long-distance running for college teachers," *Disability and Rehabilitation: Assistive Technology,* vol. 21, no. 3, p. 44–63, 2026.

[3] E. M. F. v. Sluijs, U. Ekelund, I. C. Silva, R. Guthold, A. Ha, D. Lubans, A. L. Oyeyemi, D. Ding and P. T. Katzmarzyk, "Physical activity behaviours in adolescence: current evidence and opportunities for intervention," *The Lancet Journal,* vol. 39, no. 8, pp. 429-442, 2021.

[4] M. Maashi, A. Alhefdhi, F. Alanazi and M. Rizwanullah, "RETRACTED ARTICLE: Enhancing communication for people with hearing disabilities through robust sign language recognition using deep learning and the internet of things," *Disability and Rehabilitation: Assistive Technology,* vol. 21, no. 3, p. 242–260, 2026.

[5] S. J. Lee, J. Poon, A. Jindarojanakul, C.-C. Huang, O. Viera, C. W. Cheong and J. D. Lee, "Artificial intelligence in dentistry: Exploring emerging applications and future prospects," *Journal of Dentistry,* vol. 155, no. 1, p. 105648, 2025.

[6] J. Liu, L. Liu, Y. Wu, Z. Wang and X. Li, "RETRACTED ARTICLE: MRI feature engineering and SVM framework for schizophrenia recognition," *Disability and Rehabilitation: Assistive Technology,* vol. 21, no. 3, p. 327–351, 2026.

[7] M. Shobanke, M. Bhatt and E. Shittu, "Advancements and future outlook of Artificial Intelligence in energy and climate change modeling," *Advances in Applied Energy,* vol. 17, no. 1, p. 100211, 2025.

[8] R. Zhang, "RETRACTED ARTICLE: Personalised sports rehabilitation analysis using a fitness enhanced model based on big data and deep learning," *Disability and Rehabilitation: Assistive Technology,* vol. 21, no. 3, p. 261–274, 2026.

[9] S. Komisarow and S. W. Hemelt, "School-Based Health Care and Absenteeism: Evidence from Telemedicine," *Education Finance and Policy,* vol. 19, no. 2, p. 252–282, 2024.

[10] D. Zhang, L. Zhou, J. Tao, T. Zhu and G. (. Gao, "KETCH: A Knowledge-Enhanced Transformer-Based Approach to Suicidal Ideation Detection from Social Media Content," *Information Systems Research,* vol. 36, no. 1, pp. 631-646, 2025.

[11] S. Liu, Y. Zhang, L. Zhao and Z. Liu, "Academic stress detection based on multisource data: a systematic review from 2012 to 2024," *Interactive Learning Environments,* vol. 33, no. 3, pp. 1823-1849, 2025.

[12] A. T. Turki, A. Mussetti, K. M. Scheidler, J. Kehrmann, P. Crivello, R. Hosch, F. Nensa and S. Polizzi, "Artificial Intelligence in Hematopoietic Stem Cell Transplantation Care and Complication Management," *Acta Haematol,* vol. 148, no. 5, p. 527–545, 2025.

[13] A. Razi, S. Kim, A. Alsoubai, G. Stringhini, T. Solorio, M. D. Choudhury and P. J. Wisniewski, "A Human-Centered Systematic Literature Review of the Computational Approaches for Online Sexual Risk Detection," *Proceedings of the ACM on Human-Computer Interaction,* vol. 46, no. 5, pp. 1-38, 2021.

[14] S. Siyoto and M. A. Sodik, Dasar metodologi penelitian, Yogyakarta: Literasi Media Publishing, 2015.

[15] E. Y. Sutopo and A. Slamet, Statistik inferensial, Yogyakarta: Andi, 2017.

[16] Sudjana, Metoda statistika, Bandung: PT Tarsito, 2001.

[17] A. Suryana, Metode penelitian eksperimen, Bandung: Alfabeta, 2020.

[18] N. S. Sukmadinata, Metode penelitian pendidikan, Bandung: PT Remaja Rosdakarya, 2015.

[19] M. G. Isnawan, Kuasi-eksperimen, Lombok Tengah: Nashir Al-Kutub Indonesia, 2020.

[20] Sugiyono, Metode penelitian kuantitatif, kualitatif, dan R&D, Bandung: Alfabeta, 2022.

[21] S. Siregar, Statistik parametrik untuk penelitian kuantitatif: Dilengkapi dengan perhitungan manual dan aplikasi SPSS versi 17, Jakarta: Bumi Aksara, 2013.

[22] N. Dantes, Desain eksperimen dan analisis data, Depok: Rajawali Pers, 2017.

[23] X. Cen, "RETRACTED ARTICLE: Multimodal deep learning methods for speech and language rehabilitation: a cross-sectional observational study," *Disability and Rehabilitation: Assistive Technology,* vol. 21, no. 3, p. 165–177, 2026.

[24] D. Xuanfeng, "RETRACTED ARTICLE: Gesture recognition and response system for special education using computer vision and human–computer interaction technology," *Disability and Rehabilitation: Assistive Technology,* vol. 21, no. 3, pp. 12-29, 2026.

[25] T. Kim and H. Song, ""I Believe AI Can Learn from the Error. Or Can It Not?": The Effects of Implicit Theories on Trust Repair of the Intelligent Agent," *International Journal of Social Robotics,* vol. 15, no. 1, p. 115–128, 2023.

[26] X. Li, "Machine learning-based risk warning for adolescents' prosocial behavior," *Applied Psychology: Health and Well-Being,* vol. 18, no. 3, p. e70156, 2026.

[27] A. C. Timmons, J. B. Duong and T. Chaspari, "A Call to Action on Assessing and Mitigating Bias in Artificial Intelligence Applications for Mental Health," *Perspectives on Psychological Science,* vol. 18, no. 5, pp. 29-45, 2022.

[28] R. Tripathi and M. R. Alam, "Multimodal AI in Suicide Risk Prediction: Behavioral, Physiological, and Digital Biomarkers," *Multimodal AI in Healthcare,* vol. 11, no. 2, p. 415–446, 2026.

[29] T. H. Nassif, A. B. Adler, R. D. Atkinson, K. H. Bentley, R. E. Bonomi, J. D'Olympia, E. R. Edwards, R. G. Fortgang, B. A. Gaudiano, J. C. Geraci, M. Goodman, D. Herman, E. M. Kleiman, S. Maggio and A. Mc, "SAFEGUARD: Transforming Military Suicide Prevention Through Predictive Analytics and Targeted Interventions," *Current Psychiatry Reports,* vol. 28, no. 1, p. 16, 2026.

[30] Samsinas, Sapirudin, H. Widoyo and Marzuki, "Early Intervention Strategies for Self-Harm Prevention Using AI-Driven Behavior Tracking in Teenagers," *International Journal of Research in Counseling,* vol. 4, no. 1, p. 67–76, 2025.

[31] Y. Yang, J. Huang and R. An, "AI in Social Work Research," *Artificial Intelligence in Social Work,* vol. 14, no. 1, p. 479–515, 2026.

[32] M. S. Sidhu and K. K. Sidhu, "AI Strategies for Handling Disciplinary and Cyber Bullying in Schools," *Current and Future Trends on AI Applications,* vol. 22, no. 1, p. 207–229, 2022.

1933_Similarity

ORIGINALITY REPORT

9%	5%	6%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to FAKULTAS ILMU KOMPUTER Student Paper	2%
2	www.mdpi.com Internet Source	1%
3	ejournal.uinmybatusangkar.ac.id Internet Source	1%
4	jurnal.polibatam.ac.id Internet Source	<1%
5	Diana Salas-Gómez, Ángel Denche-Zamorano, Sara Chimento-Díaz, Sabina Barrios-Fernandez. "Leisure-Time Physical Activity Frequency and Emotional/Behavioral Difficulties: a Graded Pattern in Children and Adolescents from a National Health Survey", Child Indicators Research, 2026 Publication	<1%
6	cahaya-ic.com Internet Source	<1%
7	Tuba Mutluer, Ipek Ege Gurel Ficicioglu, Melike Uysal, Sultan Ekinci et al. "A systematic review of artificial intelligence in child and adolescent interventions: from psychotherapy to developmental support", European Child & Adolescent Psychiatry, 2026 Publication	<1%
8	Olive K. L. Woo. "Artificial Intelligence in Cognitive Behavioural Therapy - A Guide for Mental Health Professionals", CRC Press, 2025 Publication	<1%
9	Daniels-Davis, Eugenia Irene. "Exploring the Attitudes, Perspectives, and Experiences of Mississippi Dental Hygienists on the Use of Artificial Intelligence (AI) in Dentistry and Related Ethical Concerns.", The University of Mississippi Medical Center Publication	<1%
10	Submitted to vcibukidnon Student Paper	<1%
11	ijere.iaescore.com Internet Source	<1%

12	doczz.net Internet Source	<1 %
13	Saul Hillman, Katharine Anderson, Richard Cross. "Screening for Trauma and Adverse Experiences in Looked-After Children Using the TALE Tool", Springer Science and Business Media LLC, 2026 Publication	<1 %
14	archive.org Internet Source	<1 %
15	www.informatica.si Internet Source	<1 %
16	journal-gehu.com Internet Source	<1 %
17	lssj.sjnl.la Internet Source	<1 %
18	Zihao Chen, Jinyi Zeng, Jinqian Liao, Cheng Guo. "The Association between Academic Stress and Problematic Internet Use among Adolescents: A Three-Level Meta-Analysis", Journal of Youth and Adolescence, 2025 Publication	<1 %
19	etheses.uin-malang.ac.id Internet Source	<1 %
20	mospace.umsystem.edu Internet Source	<1 %
21	repository.ubaya.ac.id Internet Source	<1 %
22	Ade Gafar Abdullah, Vina Adriany, Cep Ubad Abdullah. "Borderless Education as a Challenge in the 5.0 Society", CRC Press, 2020 Publication	<1 %
23	El-Sayed Atlam, M. Rokaya, M. Masud, H. Meshref, Rakan Alotaibi, Abdulqader M. Almars, Mohammed Assiri, Ibrahim Gad. "Explainable artificial intelligence systems for predicting mental health problems in autistics", Alexandria Engineering Journal, 2025 Publication	<1 %
24	Kallai Selvi Selvaraja, Nasrudin Bin Subhi, Muhammad Ajib Bin Abd Razak, Noor Rafizah Aminah Binti Aziz et al. "Development of the Healing Minds Through Play Prototype: A CBT–DBT Psychoeducational Snakes and Ladders Game to Enhance Early Engagement in Adolescent Counselling Sessions in Malaysia", Springer Science and Business Media LLC, 2026 Publication	<1 %

25	Khairuni Ulfayasha, Risydah Fadilah, Yudistira Fauzy Indrawan. "Living with Loss: Emotional Maturity and Resilience among Internal Migrant Students in Medan, Indonesia, Experiencing Paternal Loss due to Death", IJIP : Indonesian Journal of Islamic Psychology, 2025 Publication	<1 %
26	Sicen Shen, Ping Zhang, Yuchang Jin, Junxiu An. "How mindsets shape trust in AI: A dual path mechanism of self-efficacy and risk perception", Computers in Human Behavior: Artificial Humans, 2026 Publication	<1 %
27	iarjset.com Internet Source	<1 %
28	ichgcp.net Internet Source	<1 %
29	jurnal.permapendis-sumut.org Internet Source	<1 %
30	mcgill.scholaris.ca Internet Source	<1 %
31	www.blueprintsprograms.org Internet Source	<1 %
32	www.frontiersin.org Internet Source	<1 %
33	www.journals.latticescipub.com Internet Source	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		